

WHAT IS CLAIMED IS:

1. A computer program product tangibly embodied in an information carrier,
5 the computer program product including instructions that when executed generate a
graphical user interface on a display device for using a computer to display and modify a
data analysis process, the graphical user interface comprising:

a process list display for:

displaying identifications of data analysis processes, and

10 receiving an entry of an identification of a data analysis process; and

a data analysis display for:

displaying a representation of each sub-process included in the data
analysis process identified by the received entry, and

displaying a connection between each displayed sub-process,

15 wherein the data analysis display is operable to display:

a data mining sub-process for creating a data attribute by
performing an analytical process on data from an analytical processing
data source, and

one or more of sub-processes of (1) an extraction sub-process for
20 extracting data from a data source, (2) a transformation sub-process for
transforming the extracted data from a data format used by the data source
to a data format used for analytical processing, (3) a loading sub-process
for loading data into the data source used for analytical processing, and (4)
a deployment sub-process for storing the created data attribute.

25 2. The computer program product of claim 1 wherein the deployment sub-
process stores the created data attribute in one of the data source, a transactional data
store other than the data source, or a second analytical data store used for analytical
processing.

3. The computer program product of claim 1 wherein each type of the sub-processes displayed in the data analysis process display is represented by a different shape than shapes representing other types of sub-processes displayed in the analysis sub-process display.

5

4. The computer program product of claim 1 wherein the graphical user interface further comprises controls for adding types of sub-processes to the data analysis process displayed in the data analysis display.

10

5. The computer program product of claim 4 wherein the controls comprise one or more of a control for adding an extraction sub-process, a control for adding a load sub-process, a control for adding an analysis sub-process, and a control for adding a deployment sub-process.

15

6. The computer program product of claim 1 wherein the graphical user interface further comprises a control for displaying information about status of the data analysis process.

7. A computer program product tangibly embodied in an information carrier, the computer program product including instructions that when executed generate a graphical user interface on a display device for using a computer to define a data analysis process, the graphical user interface comprising:

a sub-processes display for:

25

receiving an entry of an identification of which of sub-processes of (1) an extraction sub-process for extracting data from a data source, (2) a transformation sub-process for transforming the extracted data from a data format used by the data source to a data format used for analytical processing, (3) a loading sub-process for loading data into a data source that is used for analytical processing, (4) a data mining sub-process for creating a data attribute by performing an analytical process on data from the analytical processing data source, and (5) an deployment sub-process for storing the created data attribute, and

30

receiving an entry identifying a computer program to be associated with each of the identified sub-processes such that the execution of the computer program causes the identified sub-process to be performed; and a common data display for receiving an entry of selected meta-data elements to be used in the data analysis process wherein each meta-data element is associated with a corresponding data element in the data source and with a corresponding data element in the analytical processing data source.

8. The computer program product of claim 7 wherein:
the data source is a transactional data source, and
the deployment sub-process stores the created data attribute in the transactional data source.

9. The computer program product of claim 8 wherein the deployment sub-process stores the created data attribute in one of the data source, a second transactional data store other than the transactional data source, or a second analytical data store other than the analytical data used for the data mining sub-process.

10. The computer program product of claim 7 wherein graphical user interface is operable to receive an input defining how a particular error is to be processed during the data analysis process.

11. The computer program product of claim 7 wherein the graphical user interface is operable to receive an input identifying a computing device or a component of a computing device to be used during the execution of one of the identified sub-processes.

12. The computer program product of claim 7 wherein the graphical user interface is operable to receive an input identifying an order in which each of the identified sub-processes are to be performed.

13. The computer program product of claim 7 wherein the graphical user interface is operable to receive an input identifying when the data analysis process is to be initiated.

5 14. A computer-implemented method for receiving information from a user for use in a data analysis process, the method comprising:

receiving an input identifying a data analysis process;

receiving sub-process inputs, each sub-process input identifying a sub-process associated with the data analysis process, wherein:

10 at least one of the identified sub-processes is (1) an extraction sub-process for extracting data from a transactional data source, (2) a transformation sub-process for transforming data extracted from the transactional data source from a data format used by the transactional data source to a data format used for analytical processing, (3) a loading sub-process for loading data into an analytical data source that is used for analytical
15 processing, or (4) a data mining sub-process for creating a data attribute by performing an analytical process on data from the analytical processing data source, and

at least one of the identified sub-processes is a deployment sub-process for storing a data attribute created in another of the identified sub-processes; and

20 storing the input identifying the data analysis process in association with the inputs identifying the multiple sub-processes for use in the data analysis process.

15. The method of claim 14 wherein:

the data source is a transactional data source, and

25 the deployment sub-process stores the created data attribute in the transactional data source.

16. The method of claim 15 wherein the deployment sub-process stores the created data attribute in one of the data source, a second transactional data store other than the transactional data source, or a second analytical data store used for analytical
30 processing.

17. The method of claim 14 wherein one of the sub-process inputs comprises a sub-process input identifying a computer program that causes the identified sub-process to be performed.

5 18. The method of claim 14 further comprising receiving inputs of meta-data elements to be used in the data analysis process wherein each meta-data element is associated with 1) a corresponding data element in the transactional data source, 2) a corresponding data element in the analytical process data source, or 3) both a corresponding data element in one of the transactional data sources and a corresponding
10 data element in the analytical process data source.

19. The method of claim 14 wherein each of the multiple sub-processes use a common message format.

15 20. The method of claim 14 further comprising receiving an input defining how a particular error is to be processed during the data analysis process.

21. The method of claim 14 further comprising receiving an input identifying a computing device or a component of a computing device to be used during the
20 execution of one of the multiple sub-processes.

22. The method of claim 14 further comprising receiving an input identifying an order in which the multiple sub-processes are to be performed.

25 23. The method of claim 14 further comprising receiving an input identifying when the data analysis process is to be initiated.

24. The method of claim 14 wherein the at least one of the identified sub-processes is a deployment sub-process comprises a first deployment sub-process for
30 storing a data attribute created in another of the identified sub-processes in a first data

store and a second deployment sub-process for storing the data attribute in a second data store.

25. The method of claim 24 wherein the first data store is the same as the
5 second data store.

26. The method of claim 24 wherein the first data store is different from the second data store.

10 27. The method of claim 26 wherein the first data store comprises a transactional data store and the second data store comprises an analytical data store.

28. A computer program product tangibly embodied in an information carrier, the computer program product including instructions that, when executed, receive
15 information from a user for use in a data analysis process, and computer program product being configured to

receive an input identifying a data analysis process;

receive sub-process inputs, each sub-process input identifying a sub-process associated with the data analysis process, wherein:

20 at least one of the identified sub-processes is (1) an extraction sub-process for extracting data from a transactional data source, (2) a transformation sub-process for transforming data extracted from the transactional data source from a data format used by the transactional data source to a data format used for analytical processing, (3) a loading sub-process for loading data into an analytical data source that is used for analytical
25 processing, or (4) a data mining sub-process for creating a data attribute by performing an analytical process on data from the analytical processing data source, and

at least one of the identified sub-processes is a deployment sub-process for storing a data attribute created in another of the identified sub-processes; and

store the input identifying the data analysis process in association with the inputs
30 identifying the multiple sub-processes for use in the data analysis process.

29. The computer program product of claim 28 wherein:
the data source is a transactional data source, and
the deployment sub-process stores the created data attribute in the transactional
data source.

5

30. The computer program product of claim 29 wherein the deployment sub-
process stores the created data attribute in one of the data source, a second transactional
data store other than the transactional data source, or a second analytical data store used
for analytical processing.

10

31. The computer program product of claim 28 wherein one of the sub-
process inputs comprises a sub-process input identifying a computer program that causes
the identified sub-process to be performed.

15

32. The computer program product of claim 28 further configured to receive
inputs of meta-data elements to be used in the data analysis process wherein each meta-
data element is associated with 1) a corresponding data element in the transactional data
source, 2) a corresponding data element in the analytical process data source, or 3) both a
corresponding data element in one of the transactional data sources and a corresponding
data element in the analytical process data source.

20

33. The computer program product of claim 28 wherein the at least one of the
identified sub-processes is a deployment sub-process comprises a first deployment sub-
process for storing a data attribute created in another of the identified sub-processes in a
first data store and a second deployment sub-process for storing the data attribute in a
second data store.

25

34. A system for receiving information from a user for use in a data analysis
the system comprising a processor connected to a storage device and one or more
input/output devices, wherein the processor is configured to:
receive an input identifying a data analysis process;

30

receive sub-process inputs, each sub-process input identifying a sub-process associated with the data analysis process, wherein:

at least one of the identified sub-processes is (1) an extraction sub-process for extracting data from a transactional data source, (2) a transformation sub-process for transforming data extracted from the transactional data source from a data format used by the transactional data source to a data format used for analytical processing, (3) a loading sub-process for loading data into an analytical data source that is used for analytical processing, or (4) a data mining sub-process for creating a data attribute by performing an analytical process on data from the analytical processing data source, and

at least one of the identified sub-processes is a deployment sub-process for storing a data attribute created in another of the identified sub-processes; and

store the input identifying the data analysis process in association with the inputs identifying the multiple sub-processes for use in the data analysis process.